THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 277, No. 15, Issue of April 12, pp. 13229–13236, 2002

Printed in U.S.A.

Crystal Structure of Imaginal Disc Growth Factor-2

A MEMBER OF A NEW FAMILY OF GROWTH-PROMOTING GLYCOPROTEINS FROM *DROSOPHILA MELANOGASTER**

Received for publication, November 1, 2001, and in revised form, December 21, 2001
Published, JBC Papers in Press, January 30, 2002, DOI 10.1074/jbc.M110502200

Paloma F. Varela‡, Andrea S. Llera‡§, Roy A. Mariuzza‡¶, and José Tormo‖†

From the ‡Center for Advanced Research in Biotechnology, W. M. Keck Laboratory of Structural Biology, University of Maryland Biotechnology Institute, Rockville, Maryland 20850 and ‖Centro Nacional de Biotecnologia, Universidad Autónoma de Madrid, Madrid 28049, Spain

Imaginal disc growth factor-2 (IDGF-2) is a member of a recently described family of *Drosophila melanogaster*-soluble polypeptide growth factors that promote cell proliferation in imaginal discs. Although their precise mode of action has not been established, IDGFs cooperate with insulin in stimulating the growth of imaginal disc cells. We report the crystal structure of IDGF-2 at 1.3-Å resolution. The structure shows the classical $(\beta\alpha)_8$ barrel-fold of family 18 glycosyl hydrolases, with an insertion of an $\alpha + \beta$ domain similar to that of *Serratia marcescens* chitinases A and B. However, amino acid substitutions in the consensus catalytic sequence of chitinases give IDGF-2 a less negatively charged environment in its putative ligand-binding site and preclude the nucleophilic attack mechanism of chitin hydrolysis. Particularly important is the replacement of Glu by Gln at position 132, which has been shown to abolish enzymatic activity in chitinases. Nevertheless, a modest conservation of residues that participate in oligosaccharide recognition suggests that IDGF-2 could bind carbohydrates, assuming several conformational changes to open the partially occluded binding site. Thus, IDGFs may have evolved from chitinases to acquire new functions as growth factors, interacting with cell surface glycoproteins implicated in growth-promoting processes, such as the *Drosophila* insulin receptor.

The different adult epidermal structures of the fruit fly *Drosophila melanogaster* derive from larval sheets of epithelial cells called imaginal discs. Although imaginal disc cells are dependent on soluble growth factors for their survival and proliferation, numerous attempts by homology searching and genetic analysis to identify proteins with direct mitogenic activity have been unsuccessful (1). Recently, however, protein factors with the ability to stimulate imaginal disc cell proliferation have been isolated by fractionating conditioned medium from imaginal disc cell cultures (2). These growth-promoting molecules, termed imaginal disc growth factors (IDGFs),[1] belong to a new family of glycoproteins that comprises at least five members (IDGF 1–5) having ~50% amino acid sequence identity to one another. Although no specific growth-promoting activity has been assigned to a previously described glycoprotein designated DS47 (3), its sequence homology and a similar pattern of expression suggest that DS47 represents a sixth member of the IDGF family (2).

IDGFs, which act at nanomolar concentrations, are among the first polypeptide growth factors to be identified from invertebrates (2, 4). IDGFs are expressed not only in larval imaginal discs but also throughout all developmental stages in variable patterns, from early embryo to different larval glands and tissues, as well as in adult nurse cells and oocytes (2). IDGFs are also strongly expressed in the fat body (5), in accordance with early reports showing that the fat body produces mitogenic factors (6, 7). At present, however, little is known about how IDGFs promote cell proliferation. These growth factors have been shown to act with mammalian insulin in stimulating imaginal disc cell growth through the *Drosophila* insulin receptor, suggesting a role as cofactors for *Drosophila* insulin or insulin-like molecules (2).

IDGFs present 15–25% amino acid sequence homology to family 18 glycosyl hydrolases, which includes the chitinases, enzymes that catalyze the hydrolysis of β-(1, 4)-*N*-acetyl-D-glucosamine linkages in chitin polymers of the arthropod cuticle (8). However, IDGFs have no known catalytic activity. In this respect, several proteins with sequence homology to chitinases but no detectable enzymatic activity have been described in vertebrates, indicating that the typical chitinase-like fold may be present in proteins with a wide range of biological functions other than chitin degradation (9–15). For example, mouse ECF-L has been described as a chemotactic factor for eosinophils (13), and porcine GP38K is associated with tissue remodeling (14).

Although IDGF-2 was first obtained and characterized from conditioned medium of imaginal disc cell cultures, we have found that the Schneider SL3 embryo-derived cell line produces IDGF-2 under similar growth conditions. This is in agreement with the identification of DS47, a likely member of the IDGF family, in culture supernatants of S2, also an embryo-derived cell line (3). We have purified IDGF-2 from

* This work was supported by National Institutes of Health Grant AI36900 (to R. A. M.) and a grant from the Spanish Agency for International Cooperation (AECI) (to J. T. and A. S. L.) The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

This paper is dedicated to the memory of Dr. José Tormo, colleague and friend.

The atomic coordinates and structure factors (code 1JND and 1JNE) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

† Deceased.

§ Present address: Instituto de Investigaciones Bioquímicas-Fundación Campomar, Consejo Nacional de Investigaciones Científicas y Técnicas-Universidades de Buenos Aires Av. Patricias Argentinas 435, Buenos Aires (1405), Argentina.

¶ To whom correspondence should be addressed: Center for Advanced Research in Biotechnology, 9600 Gudelsky Dr., Rockville, MD 20850. Tel.: 301-738-6243; Fax: 301-738-6255; E-mail: mariuzza@carb.nist.gov.

[1] The abbreviations used are: IDGF, imaginal disc growth factor; NAG, *N*-acetylglucosamine.

SL3 cells grown in supplemented complete medium and then allowed to condition serum-free medium. The purified glycoprotein was crystallized, and its three-dimensional structure was determined at high resolution, revealing the characteristic $(\beta\alpha)_8$ or triose-phosphate isomerase barrel-fold of family 18 glycosyl hydrolases (8). The structure enables us to explain some of the unique characteristics of IDGF-2, including the lack of chitinase activity. Certain features of the site homologous to the chitin-binding site of chitinases may be of significance for the recognition of putative ligands, such as the insulin receptor, that may account for its potent growth-promoting effects on *Drosophila* cells.

EXPERIMENTAL PROCEDURES

Production and Purification of IDGF-2—Embryo-derived Schneider SL3 *Drosophila* cells were routinely cultured in suspension (shaker flasks) at 28 °C, in Sf-900 II medium supplemented with 3% fetal bovine serum (Invitrogen). In the last passage before harvesting, cells were diluted 1:5 in *Drosophila* SFM medium (Invitrogen), without addition of fetal bovine serum. Cells were grown for 3 days in this medium and then harvested. After centrifugation, the culture supernatant was dialyzed against 20 mM Tris-HCl, pH 7.9, and concentrated 4-fold using a spiral cartridge ultrafiltration system (Amicon, Beverly, MA). The supernatant was then applied to a Q-Sepharose anion exchange column (Amersham Biosciences) equilibrated in 20 mM Tris-HCl, pH 7.9. IDGF-2 eluted at 50 mM NaCl using a linear salt gradient. The protein was further purified with a Mono Q column (Amersham Biosciences) run under similar conditions. The final yield of IDGF-2 was ~2 mg/liter culture. Protein identity was verified by N-terminal sequencing.

Crystallization and Data Collection—Two crystal forms of IDGF-2 were obtained, trigonal and orthorhombic (Table I). The trigonal form belongs to space group $P3_221$ with unit cell dimensions $a = b = 106.4$ Å, $c = 90.0$ Å, $\alpha = \beta = 90.0°$, $\gamma = 120°$. These crystals were grown by mixing aliquots of protein solution, at 10 mg/ml, with an equal volume of the reservoir solution consisting of 12% polyethylene glycol 4000, 0.1 M sodium acetate, pH 5.0, at 4 °C. The orthorhombic form belongs to space group $P2_12_12_1$ with unit cell dimensions $a = 49.5$ Å, $b = 72.2$ Å, $c = 105.8$ Å, $\alpha = \beta = \gamma = 90.0°$. These crystals were obtained at 25 °C in 25% polyethylene glycol 4000, 0.1 M Tris, pH 7.0, using IDGF-2 at 15 mg/ml. In both crystal forms, only one molecule is present in the asymmetric unit. For cryogenic data collection, crystals were harvested in a modified reservoir solution containing 20% ethylene glycol as cryoprotectant and flash-cooled by plunging into propane or liquid nitrogen. Data were collected at beamline X12B of the Brookhaven National Synchrotron Laboratory (for the trigonal form) and at beamline X11 of the Deutsche Elektronen Synchrotron (for the orthorhombic form) on ADSC quantum4 and Mar Research CCD plate detectors, respectively. Data were integrated, scaled, and merged using the HKL package (16). Data statistics are given in Table I.

Structure Determination and Refinement—The structure of IDGF-2 was determined from the trigonal crystal form by molecular replacement with truncated coordinates of *Serratia marcescens* chitinase A (17) as the search model. Normalized structure factors from 15 to 4.5 Å were used in AMoRe (18) rotation and translation functions. Model phases were improved by wARP (19). The resulting electron density map allowed unambiguous rebuilding of the molecule. The orthorhombic crystal form was readily solved by molecular replacement using partially refined coordinates from the trigonal form.

For both crystal forms, refinement was performed using CNS1.0 (20), including bulk solvent correction and overall anisotropic scaling, interspersed with iterative rounds of model rebuilding using O (21). A last round of individual anisotropic B-factor refinement was carried out using SHELX for both crystal forms (22). Only procedures that minimized both R_{cryst} and R_{free} were used. The final model includes 91.2 and 89.1% of all residues fitted to both averaged and unaveraged σ_A-weighted $2F_o - F_c$ and $F_o - F_c$ electron density maps, for the trigonal and orthorhombic forms, respectively. All regions of IDGF-2 are well ordered, with the exception of the missing loop between Val-141 and Ile-161. Both models contain residues 2 to 141 and 161 to 420; 701 and 439 solvent atoms were assigned for the trigonal and orthorhombic forms, respectively. The present $R_{cryst} = 17.6\%$ and $R_{free} = 20.2\%$ for all data ($F > 0$) between 100 and 1.3 Å for the trigonal crystal form. For the orthorhombic form, $R_{cryst} = 19.5\%$ and $R_{free} = 25.9\%$ for all data ($F > 0$) between 20 and 1.7 Å. Refinement statistics for both crystal forms are given in Table I. Atomic coordinates of the trigonal and orthorhombic forms of IDGF-2 have been deposited in the Protein Data Bank under accession codes 1JND and 1JNE, respectively (http://www.rcsb.org).

Structure superpositions were done with SHP (23). Figs. were produced with GRASP (24), MOLSCRIPT (25), or BOBSCRIPT (26) and rendered with RASTER3D (27). Sequence alignments were carried out in ClustalW at ExPASy (www.expasy.ch), and subsequently edited manually based on the known structures of chitinase A (Protein Data Bank accession code 1CNV), chitinase B (1E15), hevamine (2HVM), and Ym1 (1E9L). Sequences were retrieved from GenBank™ or SwissProt (IDGFs 1–4, AAC99417–20; IDGF-5, AAF57703.1; DS47, AAC48306; BRP39, S61550; HCGP-39, AAA16074; GP38K, AAA86482; YKL-39, AAC50597; oviduct specific glycoprotein, Q28990). Fig. 3 was drawn using ESPript (28).

RESULTS AND DISCUSSION

Overall Structure of IDGF-2: a Chitinase-like Triose-phosphate Isomerase Barrel—The structure of IDGF-2 was solved in two different crystal forms, trigonal and orthorhombic (Table I). Both forms show essentially the same structure, with a root-mean-square difference of 0.5 Å for 398 Cα atoms. The greatest deviation occurs in an external loop comprising residues 320–340 that can be attributed to differences in crystal packing. Because better resolution (1.3 Å) was achieved with the trigonal crystal, further description of the structure is based on this crystal form. The electron density maps are of high quality, as shown in Fig. 1. A carbohydrate chain can be traced attached to Asn-200 at the conserved *N*-linked glycosylation site in IDGFs (2). The visible sugar residues include two *N*-acetylglucosamines (NAGs) and two mannoses. The NAGs are linked to their neighbor by $\beta(1–4)$ linkages and a $\alpha(1–3)$ linkage binds the last Man residue.

As predicted from amino acid sequence similarity, IDGF-2 adopts the characteristic fold of family 18 glycosyl hydrolases (8), the $(\beta\alpha)_8$ barrel (residues 2–279 and 371–420), consisting of an eight-stranded parallel β-barrel ($\beta1$-$\beta8$) surrounded by eight α-helices antiparallel to the barrel ($\alpha1$-$\alpha8$) (Fig. 2). An insertion in the barrel motif between strand $\beta7$ and helix $\alpha7$ (residues 280–370) forms an additional domain with an $\alpha + \beta$-fold that is also present in *S. marcescens* chitinases A and B (17, 29), but not hevamine (30) (Fig. 2). Interestingly, this feature is common to most chitinase-like proteins described to date, although the length of the insertion varies (Fig. 3).

Comparison with Family 18 Glycosyl Hydrolases—The structures of several chitinases, as well as of proteins with chitinase-like folds but no apparent chitinase activity, have been reported, in their free forms or in complexes with ligands: chitinase A (17), chitinase B (29, 31), hevamine (30, 32, 33), concanavalin B (34), narbonin (35), endoglycosidase H (36), endoglycosidase F (37), *Coccidioides immitis* chitinase 1 (38), and the mammalian lectin Ym1 (39). Despite low sequence homology (22–25% similarity), a few conserved residues of family 18 glycosyl hydrolases, also present in IDGF-2, seem essential to maintain the barrel folding (Fig. 3; residues in *white* over *red background*) and, in the case of active chitinases, the catalytic and substrate-binding sites (33).

IDGF-2 has three *cis* peptide bonds, two of them not involving a proline residue: Gly-39–Tyr-40, Pro-295–Val-296, and Phe-394–Asp-395 (Fig. 3). The first and third are conserved in all family 18 chitinase structures and appear to be necessary for correct folding of the barrel. The aromatic residues involved in these *cis* bonds (Tyr-40 and Phe-394) seem to be important for binding of substrate in all glycosyl hydrolases with triose-phosphate isomerase barrel-folds (40). The Pro-295–Val-296 *cis* bond is located in the inserted $\alpha + \beta$ domain but is not conserved in chitinase A or B from *S. marcescens* (17).

There are two disulfide bridges in the IDGF-2 structure, which are also found in the murine chitinase-like lectin Ym1 (39). The first, between Cys-6 and Cys-33, links strand $\beta1$ with the loop between helix $\alpha1$ and strand $\beta2$. The second, formed by

TABLE I
Summary of data collection and model refinement statistics for IDGF-2

Data collection		
Space group	$P3_221$	$P2_12_12_1$
Molecules per asymmetric unit	1	1
Unit cell dimensions	$a = b = 106.4$ Å, $c = 90.0$ Å, $\alpha = \beta = \gamma = 90.0°$, $\gamma = 120°$	$a = 49.5$ Å, $b = 72.2$ Å, $c = 105.8$ Å, $\alpha = \beta = \gamma\ 90.0°$
Resolution (Å)	1.3	1.7
Total observations	758,311	303,904
Unique reflections	129,904	40,966
Completeness (%)	90.3 (53.5)[a]	98.2 (80.8)[a]
$I/\sigma<I>$	23.82 (2.88)[a]	23.27 (6.05)[a]
R_{sym}[b]	2.3 (25.7)[a]	4.3 (27.4)[a]
Refinement		
Resolution range (Å)	100–1.3	20–1.7
R_{cryst}[c]/R_{free}[d] (%)	17.6/20.2	19.5/25.9
Reflections		
Working set	127,327	36,692
R_{free} set	2,298	4,114
Number of non-hydrogen protein atoms	3,166	3,166
Number of solvent atoms	701	439
Number of heteroatoms	50	50
Root mean square deviations from ideality		
Bond lengths (Å)	0.006	0.006
Bond angles (°)	1.600	1.540
Ramachandran plot statistics[e]		
Most favored (%)	91.2	89.1
Allowed (%)	8.8	10.9

[a] Values in parentheses correspond to the highest resolution shell (1.33 to 1.30 Å for $P3_221$ and 1.74 to 1.70 Å for $P2_12_12_1$).
[b] $R_{sym}=\Sigma|I_j - <I>/\Sigma I_j$ where I_j is the intensity of an individual reflection and $<I>$ is the average intensity of that reflection.
[c] $R_{cryst}=\Sigma||F_o| - |F_c|/\Sigma|F_c||$, where F_c is the calculated structure factor.
[d] R_{free} is as for R_{cryst}, but calculated for a randomly chosen 1.8% (for $P3_221$) or 10.1% (for $P2_12_12_1$) of reflections that were omitted from the refinement.
[e] As calculated by PROCHECK (56).

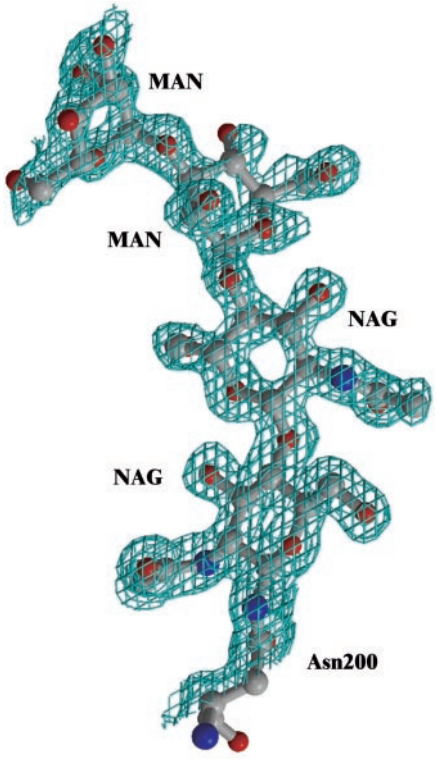

FIG. 1. **Electron density map of IDGF-2 trigonal form.** σ_A-weighted $F_o - F_c$ omit electron density map for four sugar residues linked to Asn-200. The contour level is 2σ, and the resolution is 1.3 Å.

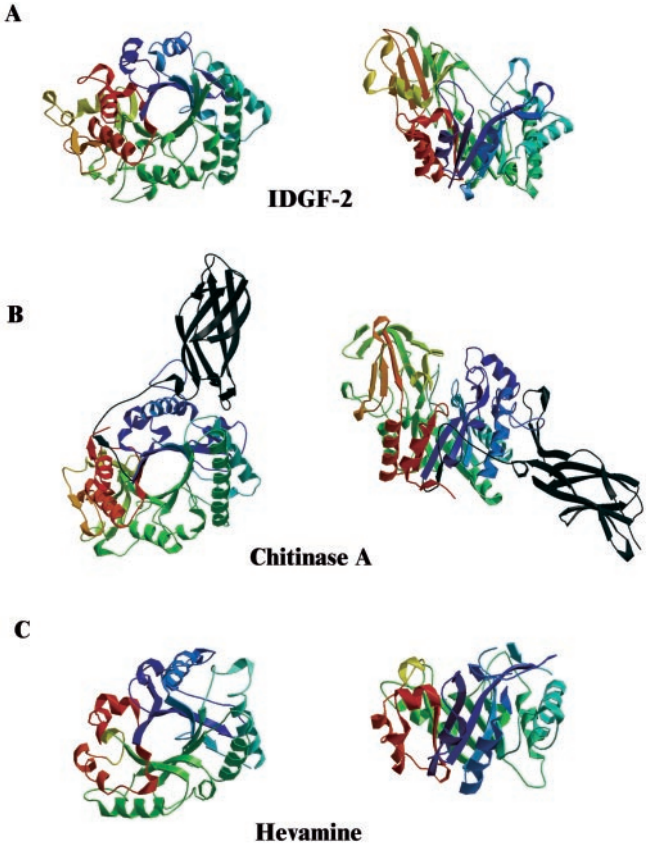

FIG. 2. **Structure of IDGF-2 and comparison with other members of the chitinase family.** Ribbons diagrams of IDGF-2 (*A*), chitinase A (*B*) (17), and hevamine (*C*) (33). The *left panels* show top views of the $(\beta\alpha)_8$ barrels, and the *right panels* show side views, after rotation by 90° around a horizontal axis. Common orientations were obtained by pairwise superpositions. The secondary elements are colored by sequence and structure homologies.

Cys-322 and Cys-405, connects the α-helix in the inserted $\alpha + \beta$ domain with the loop between strand β8 and helix α8. These Cys residues are conserved in all six IDGF family members, as well as in a human chitotriosidase (41) and in mammalian chitinase-like proteins with no chitinase activity (Fig. 3). In

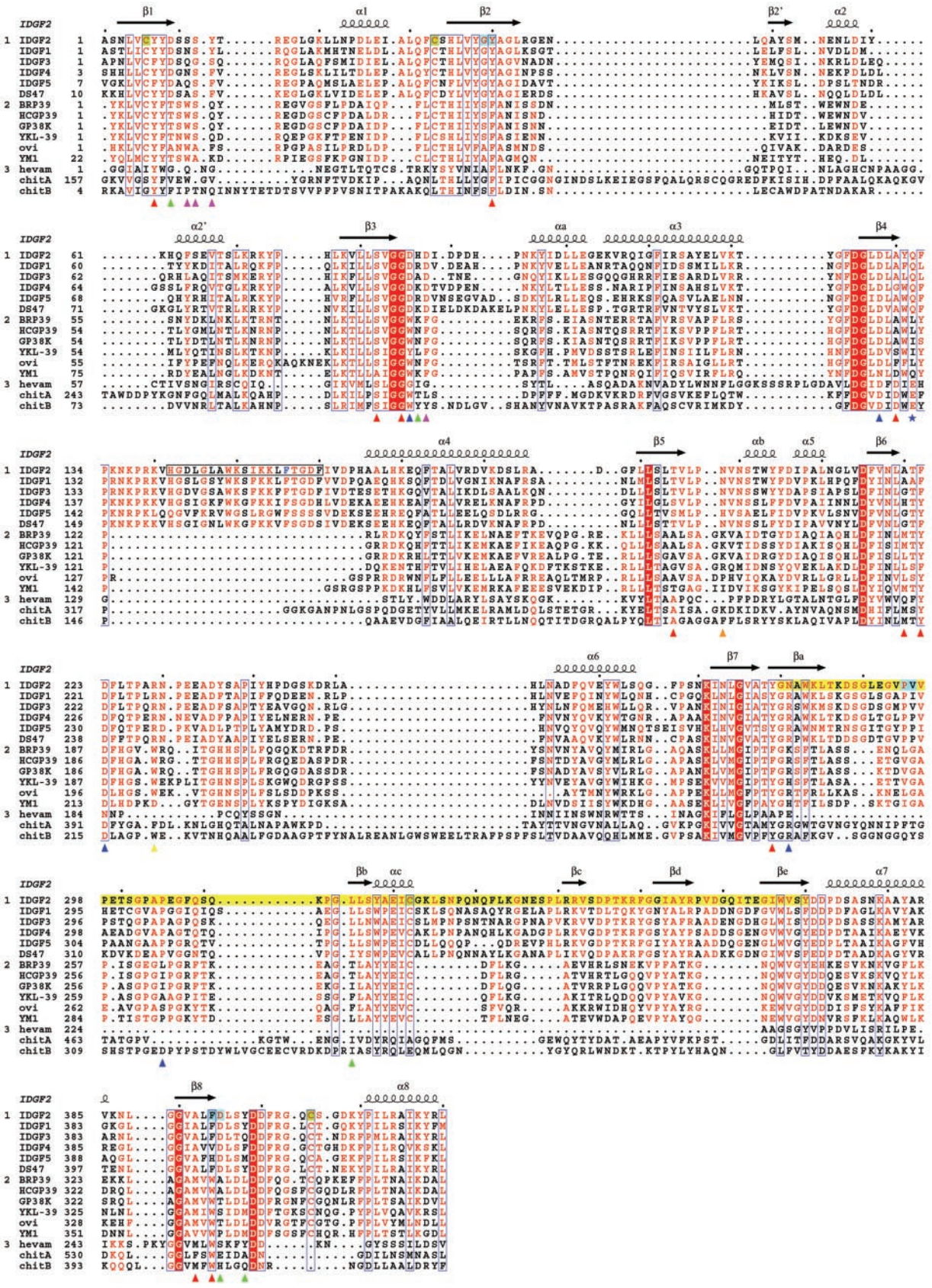

FIG. 3. **Structure-based sequence alignments of IDGFs, other chitinase-like proteins, and chitinases.** Sequences are arranged in three groups. The first contains all IDGF family members, including DS47 (2, 3). The second comprises the sequences of mammalian chitinase-like proteins with no chitinase activity that best aligned with IDGFs. These are murine mammary tumor marker BRP39 (12), human cartilage gp-39 (10), porcine smooth muscle gp38k (14), human chondrocyte YKL39 (11), porcine oviductal secretory glycoprotein (9), and murine macrophage lectin Ym1 (15). Murine eosinophil chemotactic cytokine ECF-L (13), not shown, is 97% identical to Ym1. The third group contains three family 18 chitinases with known structure: plant hevamine (30) and bacterial chitinases A (17) and B (29). Chitinase A was used as a model for solving the IDGF-2 structure by molecular replacement as it showed the highest sequence homology. In some multidomain proteins, sequences were

contrast, the Cys residues involved in disulfide bond formation in IDGF-2 are not present in family 18 glycosyl hydrolases from plants or bacteria.

In $(\beta\alpha)_8$ barrel enzymes, loops between β strand C termini and the N-terminal portion of α-helices (loops $\beta X\alpha X$, X being the number of the strand in the barrel) are usually involved in interactions with substrate and in catalytic activity. These loops, although fairly variable in length and amino acid composition, conserve some key residues responsible for catalysis and substrate binding (33). This is also true for chitinase-like proteins with no apparent enzymatic activity, as seen in Fig. 3. Loop $\beta 4\alpha 4$ (residues 131–165) is distinctly long in the IDGF family. Although neither of our two crystal forms shows electron density between Val-141 and Ile-161, the location of these boundary residues on the surface of the IDGF-2 structure suggests that the $\beta 4\alpha 4$ loop should be fully exposed to solvent. Furthermore, this loop appears to have undergone proteolytic cleavage. Thus, SDS-PAGE analysis of purified IDGF-2, as well as of redissolved crystals, revealed only a faint band at 47 kDa (the expected molecular mass of the intact protein), along with two prominent bands at 18 and 29 kDa that together represented >90% of the total protein (data not shown). N-terminal sequencing of the 18-kDa band gave the sequence ASNLV*X*YYDSS*X*YTREGLGK, corresponding to the predicted N terminus following removal of the signal peptide during secretion of mature IDGF-2 to the external medium. The 29-kDa band yielded the sequence TGDFIVDPHAALHKEQ, implying a nick between Phe-156 and Thr-157. The lack of electron density for most residues of the 131–165 loop may reflect disorder resulting from this cleavage. It is also possible that at least some of residues 142–160 have been excised in our preparations. Importantly, cleaved and uncleaved IDGF-2 display comparable activity in imaginal disc cell proliferation assays, demonstrating that an intact 131–145 loop is not essential for IDGF-2 function.[2]

Analysis of the IDGF-2 Region Homologous to the Catalytic Site of Chitinases—The active site of chitinases is situated in a cleft formed by the C termini of the β-strands and loops $\beta X\alpha X$, which presents aromatic and negatively charged residues to the substrate (Fig. 4*B*). A Glu residue (*e.g.* Glu-315 in chitinase A) is the key proton donor during hydrolysis of the glycosidic bond (42, 43) and is usually imbedded in a hydrophobic environment that contributes to substrate binding, with other negatively charged residues facilitating the nucleophilic attack of the catalytic Glu. In IDGF-2 and other IDGF family members, the corresponding residue is Gln-132 (Fig. 3, *blue star*; Fig. 4*A*). This substitution is seen in other chitinase-like proteins with no apparent enzymatic activity (Fig. 3) (13, 34, 44, 45), for example Ym1 (Fig. 4*C*), and has been shown to abolish enzymatic activity in chitinases (38, 42, 45), as well as in endoglycosidase H, a family 18 glycosyl hydrolase (46). Interestingly, the side chain of Gln-132 in IDGF-2 is oriented differently from that of the catalytic Glu in chitinases, as it is displaced away from the center of the putative ligand-binding cavity by the side chain of Leu-196 and by Pro-197 (Fig. 4, *A* and *B*). This conformation of Gln may be exclusive to IDGFs, as no other chitinase or chitinase-like protein appears to conserve Leu and Pro at

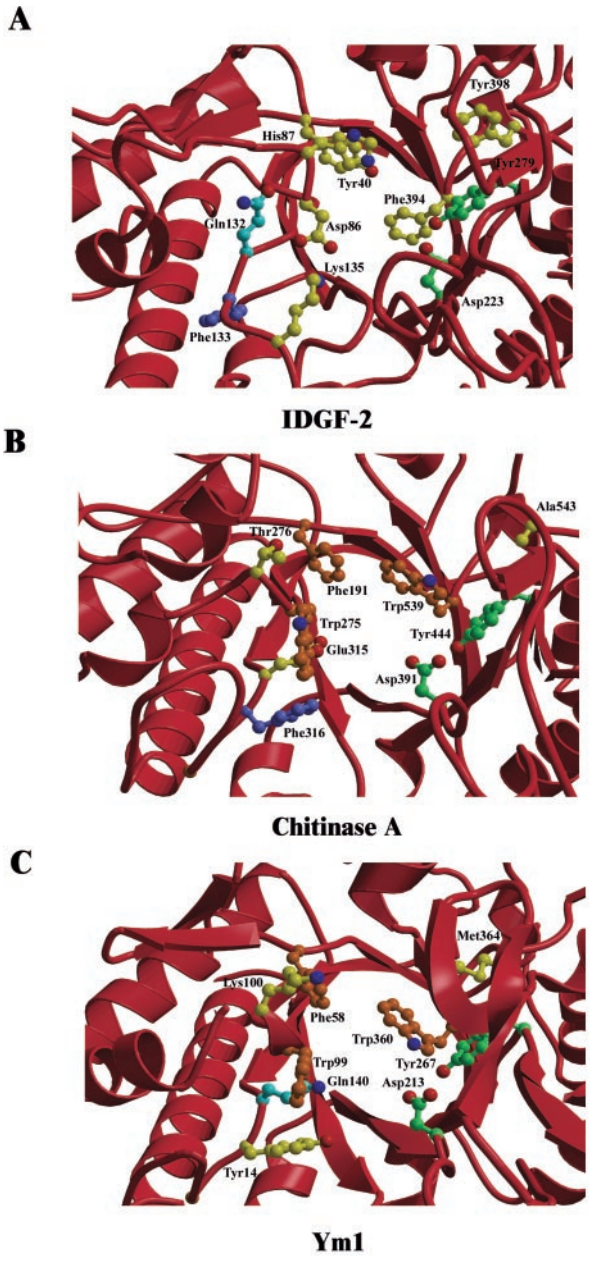

FIG. 4. **Comparison of the binding sites of IDGF-2, a chitinase, and a chitinase-like protein.** *A*, putative ligand-binding site of IDGF-2. *B*, same view of the active site of chitinase A (17). *C*, saccharide binding site of Ym1 (39). Residues conserved in all three structures are *green*; those not conserved in any of the structures are *yellow*. Residues conserved between IDGF-2 and chitinase A are *lilac*, residues conserved between IDGF-2 and Ym1 are *cyan*, and residues conserved between chitinase A and Ym1 are *orange*.

[2] P. J. Bryant, personal communication.

truncated to display only the significant domain. Residue numbers are those of the mature proteins, except in the case of chitinase A and Ym1 where PDB numbering takes into account the signal peptide sequence. IDGF-2 secondary structure elements are depicted on top as assigned by DSSP (57). *Squiggles* represent helices, and *arrows* represent β strands. Helices and β strands that are part of the triose-phosphate isomerase barrel are *numbered* (including an apostrophe if they are discontinuous), and secondary elements outside the barrel are labeled with *letters*. Every tenth residue of IDGF-2 is marked with a *dot* over its sequence. *White* characters on a *red background* show residues strictly conserved in all groups. Residues well conserved within each group are in *red* characters, and the rest are in *black*. A *blue frame* denotes similarity across groups. Cys residues that form disulfide bridges are in a *light green background*. Residues in a *cyan background* are involved in *cis* peptide bonds. The missing loop (residues 142–160) is boxed in *black*, with a *blue F* indicating where the nick seen by SDS-PAGE and N-terminal sequencing is located. The $\alpha + \beta$ domain insertion is marked in *yellow*. *Arrowheads* mark residues involved in carbohydrate binding in chitinases (29, 31). *Purple* indicates subsite −3, *green* subsite −2, *red* subsite −1, *blue* subsite +1, *yellow* subsite +2, and *orange* subsite +3. The position equivalent to the catalytic Glu is marked with a *blue star*.

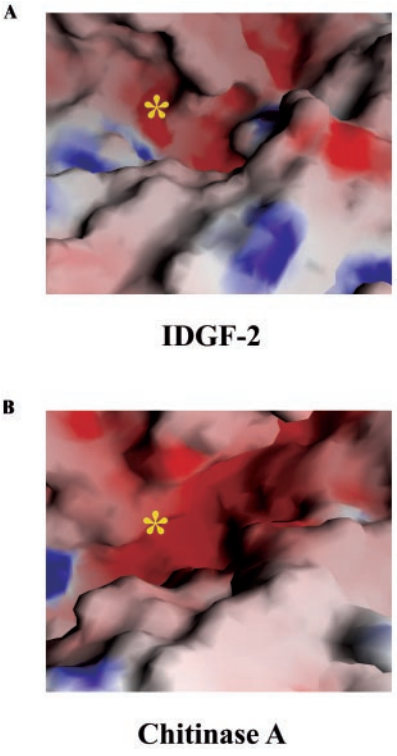

FIG. 5. **Comparative surface analysis of the binding sites of IDGF-2 and chitinase A.** Electrostatic surface potentials for IDGF-2 (*A*) and chitinase A (*B*) (17) were calculated using GRASP (24). Solvent-accessible surfaces are colored according to electrostatic potential, with positively charged residues in *blue* and negatively charged residues in *red*. The positions of residues Gln-132 of IDGF-2 (*A*) and the catalytic Glu-315 of chitinase A (*B*) are marked by *asterisks*.

these positions (Fig. 3). In fact, Gln-120 of the chitinase-like lectin Ym1, which corresponds to Gln-132 of IDGF-2, is oriented similarly to the catalytic Glu of chitinases (Fig. 4, *A* and *C*).

In terms of electrostatic surface potential, comparison of the active sites of chitinases and chitinase-like lectins with the equivalent region in IDGF-2 shows that, overall, the IDGF-2 cleft is less negatively charged (Fig. 5), due to several changes in otherwise conserved acidic residues. Apart from the substitution of Glu for Gln at position 132, there is the replacement of Asp by Ala at position 130 (Fig. 3). In chitinases, this Asp is believed to be involved in determining the physicochemical characteristics of the proton donor during polysaccharide hydrolysis; in the structures of hevamine and chitinase B complexed with ligands, it fixes the catalytic Glu in position with a hydrogen bond (31, 33). However, in IDGF-2 the Ala substitution prevents this hydrogen bond-mediated mechanism. Interestingly, this substitution creates a cavity similar to that formed by a conformational change in Asp-142 upon sugar binding to chitinase B (31). In this conformation, contrary to what is seen in apo-chitinase B for equivalent residues, the phenolic group of Tyr-7 is within hydrogen bonding distance of Asp-128, whereas Ser-82 fills the aforementioned cavity and contributes to strengthening this hydrogen bond.

In chitinases, enzymatic activity comprises two distinct processes: chitin binding and hydrolysis. These activities are separable to the extent that several proteins with chitin binding, but no chitinolytic, activity have been identified (13, 41). The active sites of chitinases contain a number of aromatic residues that are implicated in carbohydrate binding (Fig. 4*B*) (31, 33). As chitin is a polymeric sugar, different regions, or subsites, in the binding cleft contact different sugar subunits and have been arbitrarily designated subsites −4 to +3 (from the nonreducing to reducing end), according to their position with respect to the sugars that undergo hydrolysis (the scissile bond is between −1 and +1) (33). Residues situated outside the binding cleft may also contribute to substrate binding, and several chitinases bear an additional chitin-binding domain that extends the substrate-binding region and helps ensure specificity (17, 29).

It has been proposed that IDGF-2 may have evolved from chitinases to acquire new properties as a lectin that enable it to bind to oligosaccharides on cell surface glycoproteins implicated in growth-promoting processes (2). Consistent with this hypothesis, chitin is composed of NAG residues, and NAG is the proximal sugar in all *N*-linked oligosaccharides. Although no biochemical evidence currently exists for IDGF-2 binding to carbohydrates,[2] certain structural features suggest that this growth factor could potentially accommodate oligosaccharide ligands interacting through hydrophobic and polar contacts. At least a portion of the putative binding cleft appears accessible to sugar molecules, although it is considerably narrower than those of chitinases (see below). Several hydrophobic residues of IDGF-2, including Tyr-7, Tyr-40, Phe-222, Tyr-279, and Phe-394, are highly conserved in family 18 glycosyl hydrolases and in other IDGFs. In the IDGF-2 structure, these residues are located in the β strands (Tyr-7, Tyr-40, Tyr-279, Phe-394) or loops (Phe-222) of the triose-phosphate isomerase barrel domain around the putative saccharide-binding pocket (Fig. 4*A*). Moreover, a few acidic residues involved in binding and/or hydrolysis of substrates by chitinases are conserved in the IDGF family, in particular Asp-128 and Asp-223 (Fig. 3). Another conserved acidic residue, Asp-125, participates in a hydrogen bonding network at the core of the barrel. However, its effect on substrate binding may be negligible, as mutants of this residue retain full enzymatic activity in chitinases (46). Overall, however, IDGFs display only a modest conservation of residues that participate in oligosaccharide recognition by chitinases (31, 32) and chitinase-like lectins (39). Furthermore, a closer inspection of the IDGF-2 structure indicates that this growth factor lacks a proper configuration of residues for binding saccharides in the same way as these latter proteins.

Superposition of the IDGF-2 structure onto those of the allosamidin-hevamine (32) and *N*,*N*′,*N*″-triacetylchitotriose (NAG_3)-hevamine (33) complexes revealed steric clashes between the bound sugars and a number of IDGF-2 residues, including His-87, Tyr-279, Lys-312, and Phe-394. These residues, as well as Tyr-40, Asp-86, and Lys-135, contribute to a substantial occlusion of the putative IDGF-2 binding site (Fig. 4*A*) compared with those of chitinase A (Fig. 4*B*) and Ym1 (Fig. 4*C*). As shown in Fig. 6, numerous clashes between IDGF-2 residues (Asp-86, His-87, Lys-135, Asn-136, Arg-229) and sugar atoms were also noted upon superposition of IDGF-2 onto chitinase B in complex with NAG_5 (31). Consistent with this analysis, attempts to co-crystallize IDGF-2 with allosamidin, NAG_2, or NAG_3 were unsuccessful, even at millimolar concentrations of the ligand, yielding only crystals of the uncomplexed growth factor. Indeed, as discussed below, significant conformational changes in a number of IDGF-2 residues would be required to permit IDGF-2 to accommodate oligosaccharide ligands in a manner analogous to chitinases or the heparin-binding lectin Ym1.

In the region of IDGF-2 homologous to subsite −1, Phe-394 replaces a Trp residue in chitinases and chitinase-like proteins (Fig. 3). Although this represents a conservative substitution, the Phe-394 side chain is rotated about the Cα-Cβ axis by 130° relative to the position of the corresponding Trp-539 side chain of chitinase A, partially occluding the putative binding cleft (Fig. 4, *A* and *B*). This conformation prevents hydrophobic contacts with sugar at subsite −1. Another conservative substitution of a

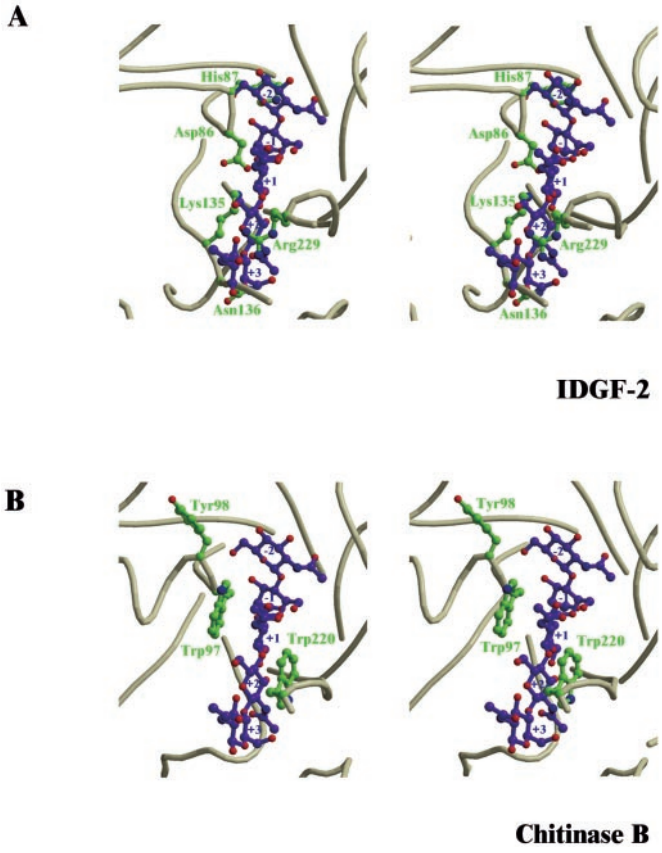

FIG. 6. **Stereo images of *N*-acetylglucosamine pentamer complexes.** *A*, model of IDGF-2 in complex with NAG_5 built by superimposing IDGF-2 onto the chitinase B-NAG_5 complex (*B*) (31). Residues of IDGF-2 that clash with the sugar in the hypothetical IDGF-2-NAG_5 complex are depicted. Indicated in *B* are chitinase B residues Trp-97, Tyr-98, and Trp-220 that correspond to IDGF-2 residues Asp-86, His-87, and Arg-229, respectively, in structure-based sequence alignments (Fig. 3). IDGF-2 residues Lys-135 and Asn-136 are part of an extended loop not found in chitinase B. Carbon atoms of the side chains are *green*, nitrogen atoms *blue*, and oxygen atoms *red*. Relevant stretches of the polypeptide backbone are *cream*. Carbon atoms of the oligosaccharide are *purple*, nitrogen atoms *blue*, and oxygen atoms *red*. Labels identify amino acid side chains and the sugar subunits bound to subsites −2 to +3.

highly conserved Tyr residue by Phe-222 would preclude a hydrogen bond to sugar, as in the NAG_5-chitinase B complex (31), but would allow a stacking interaction with sugar in subsite −1. Also contributing to a narrowing of subsite −1 is Tyr-40, which replaces a conserved Phe residue. At subsite +1, Asp-86 substitutes for Trp-275 in chitinase A and Trp-99 in Ym1, forming a salt bridge with Lys-135 that practically closes the cleft at this point (Fig. 4). The side chain of IDGF-2 residue Tyr-398, whose conformation is stabilized by a hydrogen bond to Lys-312, would appear to block the entry of a sugar molecule at subsite −2. The corresponding residue in hevamine (Tyr-259) points out from this subsite toward solvent (30). Particularly important is IDGF-2 His-87, the side chain of which contributes to a further narrowing of subsite −2 (Figs. 4*A* and 6*A*).

Other subsites are less conserved in IDGF-2. Thus, the side chain of Arg-229, which hydrogen bonds with the main chain of Asn-136, occludes subsite +2 (Fig. 6*A*). The equivalent position in chitinase A and B is occupied by an aromatic residue (Phe-396 and Trp-220, respectively), the side chain of which stacks on a sugar ring (Fig. 6*B*). Finally, the putative ligand-binding cleft of IDGF-2 is surrounded by two loops, $\beta 4\alpha 4$-(131–165) and $\beta 5\alpha 5$-(196–199) that, along with the $\alpha + \beta$ insertion, would prevent the extension of polymeric sugars beyond subsite +3 (Fig. 6*A*). Based on sequence alignments (Fig. 3), both loops are expected to be present in all six IDGF family members.

The availability of crystal structures of hevamine and chitinase B in free and liganded forms allows a direct assessment of the magnitude of conformational changes in these chitinases associated with saccharide binding. Only minor structural rearrangements in the hevamine active site were noted in comparisons of the free enzyme with allosamidin-hevamine (32) and NAG_3-hevamine (33) complexes. In the case of chitinase B, several significant conformational changes occur upon binding a NAG_5 substrate (31). However, these are mainly associated with the catalytic mechanism (*i.e.* stabilization of an intermediate oxazolinium ion) and closure of the roof of the active site tunnel, rather than with the primary binding interaction. Indeed, oligosaccharide substrates could be readily modeled into the binding cleft of the free chitinase B structure without invoking conformational changes to relieve unfavorable contacts (29).

In contrast to chitinases, the accommodation of oligosaccharide ligands by IDGF-2 would appear to require significant rearrangements in a number of residues to open the partly blocked binding site. Although such structural changes (and their associated energetic costs) cannot, of course, be excluded on the basis of present data, the possibility also exists that IDGFs may instead reflect an evolutionary progression from sugar-binding to primarily protein-binding molecules that could retain some sugar-binding capacity. Such an evolution is well documented in the C-type lectin family, which includes *bona fide* Ca^{2+}-dependent lectins (mannose-binding protein, selectins, tunicate lectin), as well as other members that recognize proteins directly, with little or no involvement of carbohydrates (natural killer cell receptors, CD23, coagulation factors IX/X-binding protein) (47, 48).

A Possible Basis for Cooperation between IDGFs and Insulin in Promoting Cell Proliferation—Imaginal discs express an insulin receptor homologous to that of mammals (49–51) that is required for their normal growth (52). The possible involvement of lectin-like proteins in activating the insulin receptor is suggested by several lines of evidence (51, 53), including the finding that mutagenesis of one of the acceptor asparagines for *N*-linked glycosylation results in major alterations in tyrosine kinase activity and in the inability to transduce signals for glycogen or DNA synthesis (54). We speculate that the observed cooperation between IDGFs and insulin in stimulating imaginal disc cell proliferation might be explained by a requirement for the insulin receptor to interact with both proteins to achieve optimum signaling. In this regard, one mechanism by which accessory, or costimulatory, molecules enhance signal transduction was illustrated recently by the crystal structure of fibroblast growth factor bound to its receptor and heparin, in which the signaling complex is assembled around a central heparin molecule linking two growth factor ligands into a dimer that bridges two receptor chains (55). Similarly, IDGFs might stabilize the binding of *Drosophila* insulin to its receptor through a simultaneous interaction with both molecules to form a multiprotein signaling complex. However, a rigorous assessment of this (or other) hypothesis to explain the growth-promoting activity of IDGFs must await biochemical studies to define the ligand binding specificity of these novel invertebrate growth factors.

Acknowledgments—We are indebted to Peter Bryant (Dept. of Developmental and Cell Biology, University of California, Irvine) for critical reading of this manuscript and for providing valuable unpublished information regarding IDGFs. We would like to thank Anastassis Perrakis for help with wARP, Eric Sundberg for comments on the manuscript, Mark Sawicki for advice on SHELX, Jianying Yang for help with the figures, and Mark Garfield for protein sequencing. We are also

grateful to the staff of beamline X12B at the Brookhaven National Synchrotron Laboratory and of beamline X11 at the EMBL-Hamburg Outstation for assistance with data collection.

REFERENCES

1. Hipfner, D. R., and Cohen, S. M. (1999) *Bioessays* **21,** 718–720
2. Kawamura, K., Shibata, T., Saget, O., Peel, D., and Bryant, P. J. (1999) *Development* **126,** 211–219
3. Kirkpatrick, R. B., Matico, R. E., McNulty, D. E., Strickler, J. E., and Rosenberg, M. (1995) *Gene* **153,** 147–154
4. Homma, K., Matsushita, T., and Natori, S. (1996) *J. Biol. Chem.* **271,** 13770–13775
5. Bryant, P. J. (2001) *Novartis Found Symp.* **237,** 182–194
6. Davis, K. T., and Shearn, A. (1977) *Science* **196,** 438–440
7. Shearn, A., Davis, K. T., and Hersperger, E. (1978) *Dev. Biol.* **65,** 536–540
8. Henrissat, B., and Davies, G. (1997) *Curr. Opin. Struct. Biol.* **7,** 637–644
9. Buhi, W. C., Alvarez, I. M., Choi, I., Cleaver, B. D., and Simmen, F. A. (1996) *Biol. Reprod.* **55,** 1305–1314
10. Hakala, B. E., White, C., and Recklies, A. D. (1993) *J. Biol. Chem.* **268,** 25803–25810
11. Hu, B., Trinh, K., Figueira, W. F., and Price, P. A. (1996) *J. Biol. Chem.* **271,** 19415–19420
12. Morrison, B. W., and Leder, P. (1994) *Oncogene* **9,** 3417–3426
13. Owhashi, M., Arita, H., and Hayai, N. (2000) *J. Biol. Chem.* **275,** 1279–1286
14. Shackelton, L. M., Mann, D. M., and Millis, A. J. (1995) *J. Biol. Chem.* **270,** 13076–13083
15. Chang, N.-C. A., Hung, S.-I., Hwa, K.-Y., Kato, I., Chen, J.-E., Liu, C.-H., and Chang, A. C. (2001) *J. Biol. Chem.* **276,** 17497–17506
16. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
17. Perrakis, A., Tews, I., Dauter, Z., Oppenheim, A. B., Chet, I., Wilson, K. S., and Vorgias, C. E. (1994) *Structure* **2,** 1169–1180
18. Navaza, J. (1994) *Acta Crystallogr. Sect. A* **50,** 157–163
19. Perrakis, A., Ouzounis, C., and Wilson, K. S. (1997) *Folding Des.* **2,** 291–294
20. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S. Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54,** 905–921
21. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47,** 110–119
22. Sheldrick, G., and Shneider T. (1997) *Methods Enzymol.* **277,** 319–343
23. Stuart, D. I., Levine, M., Muirhead, H., and Stammers, D. K. (1979) *J. Mol. Biol.* **134,** 109–142
24. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins* **11,** 281–296
25. Kraulis, P. J. (1991) *Science* **254,** 581–582
26. Esnouf, R. M. (1997) *J. Mol. Graph. Model* **15,** 132–134
27. Merritt, E. A., and Murphy, M. E. P. (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50,** 869–873
28. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) *Bioinformatics* **15,** 305–308
29. van Aalten, D. M. F., Synstad, B., Brurberg, M. B., Hough, E., Riise, B. W., Eijsink, V. G. H., and Wierenga, R. K. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 5842–5847
30. Terwisscha van Scheltinga, A. C., Kalk, K. H., Beintema, J. J., and Dijkstra, B. W. (1994) *Structure* **2,** 1181–1189
31. van Aalten, D. M. F., Komander, D., Synstad, B., Gaseidnes, S., Peter, M. G., and Eijsink, V. G. H. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 8979–8984
32. Terwisscha van Scheltinga, A. C., Armand, S., Kalk, K. H., Isogai, A., Henrissat, B., and Dijkstra, B. W. (1995) *Biochemistry* **34,** 15619–15623
33. Terwisscha van Scheltinga, A. C., Hennig, M., and Dijkstra, B. W. (1996) *J. Mol. Biol.* **262,** 243–257
34. Hennig, M., Jansonius, J. N., Terwisscha van Scheltinga, A. C., Dijkstra, B. W., and Schlesier, B. (1995) *J. Mol. Biol.* **254,** 237–246
35. Hennig, M., Schlesier, B., Dauter, Z., Pfeffer, S., Betzel, C., Hohne, W. E., and Wilson, K. S. (1992) *FEBS Lett.* **306,** 80–84
36. Rao, V., Guan, C., and Van Roey. P. (1995) *Structure* **3,** 449–457
37. Van Roey, P., Rao, V., Plummer, T. H., Jr., and Tarentino, A. L. (1994) *Biochemistry* **33,** 13989–13996
38. Hollis, T., Monzingo, A. F., Bortone, K., Ernst, S., Cox, R., and Robertus, J. D. (2000) *Protein Sci.* **9,** 544–551
39. Sun, Y.-J., Chang, N.-C. A., Hung, S.-I., Chang, A. C., Chou, C.-C., and Hsiao, C.-D. (2001) *J. Biol. Chem.* **276,** 17507–17514
40. Jabs, A., Weiss, M. S., and Hilgenfeld, R. (1999) *J. Mol. Biol.* **286,** 291–304
41. Renkema, G. H., Boot, R. G., Au, F. L., Donker-Koopman, W. E., Strijland, A., Muijsers, A. O., Hrebicek, M., and Aerts, J. M. F. G. (1998) *Eur. J. Biochem.* **251,** 504–509
42. Lin, F.-P., Chen, H.-C., and Lin, C.-S. (1999) *IUBMB Life* **48,** 199–204
43. McCarter, J. D., and Withers, S. G. (1994) *Curr. Opin. Struct. Biol.* **4,** 885–892
44. Jin, H. M., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., Kirkpatrick, R. B., and Rosenberg, M. (1998) *Genomics* **54,** 316–322
45. Watanabe, T., Kobori, K., Miyashita, K., Fujii, T., Sakai, H., Uchida, M., and Tanaka, H. (1993) *J. Biol. Chem.* **268,** 18567–18572
46. Schmidt, B. F., Ashizawa, E., Jarnagin, A. S., Lynn, S., Noto, G., Woodhouse, L., Estell, D. A., and Lad, P. (1994) *Arch. Biochem. Biophys.* **311,** 350–353
47. Weis, W. I., Taylor, M. E., and Drickamer, K. (1998) *Immunol. Rev.* **163,** 19–34
48. Sawicki, M. W., Dimasi, N., Natarajan, K., Wang, J., Margulies, D. H., and Mariuzza, R. A. (2001) *Immunol. Rev.* **181,** 52–65
49. Garofalo, R. S., and Rosen, O. M. (1988) *Mol. Cell. Biol.* **8,** 1638–1647
50. Fernández, R., Tabarini, D., Azpiazu, N., Frasch, M., and Schlessinger, J. (1995) *EMBO J.* **14,** 3373–3384
51. Ruan, Y., Chen, C., Cao, Y., and Garofalo, R. S. (1995) *J. Biol. Chem.* **270,** 4236–4243
52. Chen, C., Jack, J., and Garofalo, R. S. (1996) *Endocrinology* **137,** 846–856
53. Marin-Hincapie, M., and Garofalo, R. S. (1995) *Endocrinology* **136,** 2357–2366
54. Leconte, I., Carpentier, J. L., and Clauser, E. (1994) *J. Biol. Chem.* **269,** 18062–18071
55. Pellegrini, L., Burke, D. F., von Delft, F., Mulloy, B., and Blundell, T. L. (2000) *Nature* **407,** 1029–1034
56. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallog.* **26,** 283–291
57. Kabsch, W., and Sander, C. (1983) *Biopolymers* **22,** 2577–2637